

Eötvös Lóránd Tudományegyetem
Pedagógiai és Pszichológiai Kar

Pszichológia Doktori Iskola
Kognitív Pszichológia Doktori Oktatási Program

Honbolygó Ferenc

A beszéd prozódiai jellemzőinek észlelése

A doktori értekezés tézisei

Témavezető: Dr. Csépe Valéria, DSc

2009

Tartalomjegyzék

1. A beszéd prozódiai jellemzőinek észlelése	1
1.1. Bevezető	1
1.2. Szakirodalmi áttekintés	2
1.3. Tézisek	6
1.4. Eredmények	6
1.4.1. 1. kísérlet	6
1.4.2. 2-4. kísérlet	8
1.4.3. 5. kísérlet	9
1.5. Megbeszélés	11
1.6. A disszertáció témájában megjelent publikációk	13
 2. The perception of prosodic features of speech	 15
2.1. Introduction	15
2.2. Literature review	16
2.3. Theses	19
2.4. Results	20
2.4.1. Experiment 1.	20
2.4.2. Experiment 2-4.	21
2.4.3. Experiment 5.	22
2.5. Discussion	24
2.6. Publications related to the topic of the dissertation	26

1.

A beszéd prozódiai jellemzőinek észlelése

1.1. Bevezető

Disszertációmban a beszéd észlelésének agyi hátterével foglalkozom. A beszéd feldolgozása során számos különböző folyamat zajlik az agyban, amelyek lehetővé teszik, hogy az akusztikai bemenetből kinyerjük a nyelvi üzenetet. Ezen folyamatok közül disszertációmban azokról esik szó, amelyek a beszéd prozódiai részével, ezen belül is a hangsúllyal kapcsolatosak. A beszéd prozódia jellemzőinek tanulmányozása viszonylag kevés figyelmet kapott a beszédészleléssel foglalkozó kutatásokban, miközben a prozódiai jellemzők a nyelvi inputnak ugyanolyan szerves részét képezik, mint a beszédhangokkal kapcsolatos információk. Ezért disszertációm alapvető célja, hogy felhívja a figyelmet a prozódiai jellemzők fontosságára.

Az értekezés központi témája a beszédészlelést kísérő agyi elektromos aktivitás tanulmányozása. Az emberi fejről elvezethető EEG jel ingerhez kötött szinkronizált megváltozása, az eseményhez kötött agyi potenciál (EKP) a kognitív idegtudomány talán legsikeresebb módszere. Segítségével viszonylag kis erőfeszítéssel betekintést nyerhetünk olyan feldolgozási folyamatokba, amelyeket a kognitív pszichológiai egyéb eszközeivel nehezen tudnánk vizsgálni. Disszertációmban bizonyítani szeretném, hogy az eseményhez kötött agyi potenciál komponensek a megfelelő kísérleti eljárások alkalmazásával lehetővé teszik azt, hogy a nyelvészeti, pszicholingvisztikai modellek által feltételezett reprezentációk működését az emberi agyban valós időben, online

módon ragadjuk meg.

Disszertációmban a hangsúly tanulmányozásában multidiszciplináris megközelítést alkalmaztam, mivel úgy vélem, hogy e komplex nyelvi jelenség megfelelő megértéséhez több elméleti keretet is figyelembe kell venni az adatok magyarázatában. Ennek megfelelően dolgozatomban szakirodalmi áttekintés részeként a hangsúly nyelvészeti, pszichológiai és idegtudományi elméleteire is kitérek.

A disszertációban megfogalmazott hipotéziseket öt empirikus kísérletben vizsgáltam. A kísérletek mindegyikében az eseményhez kötött agyi potenciálok módszerét alkalmaztam. A vizsgálatok során az EKP komponensek közül az Eltérési Negativitás komponenst, mint a preattentív változásdetekcióhoz kapcsolódó agyi elektromos változást tanulmányoztam.

1.2. Szakirodalmi áttekintés

A beszéd észlelése során alapvetően két típusba sorolható akusztikai információt kell feldolgoznunk: a beszédhangok, hangkapcsolatok és hangsorok jellemzőire vonatkozó *szegmentális* információt, és az ezekre mintegy ráépülő, a hangsorok egészét moduláló, több beszédhangon átívelő *szupraszegmentális* információt (Gósy, 2004).

A prozódiai vagy szupraszegmentális jellemzőket hagyományosan a szegmentális jellemzőkkel állítják szembe. A beszédprodukción ez a két jellemző együtt, komplex hangjelként jelenik meg, egyidőben hozzuk létre, és a beszédpercepció során összetartozóan, egyidejűleg kell feldolgoznunk őket. A két jellemzőnek a nyelvi értésben eltérő funkciói vannak. Leegyszerűsítve azt mondhatjuk, hogy míg a szegmentális jellemzők a *beszédhangokkal* kapcsolatos akusztikai információt tartalmazzák, addig a szupraszegmentális jellemzők elsősorban *tagoló* funkciót töltenek be, azaz a különböző szintű nyelvi egységek (szavak, frázisok, mondatok) elválasztását teszik lehetővé.

Disszertációm középpontjában a prozódiai jellemzők közül az egyik legvitatottabb jellemző, a *hangsúly* áll. A hangsúly természetével kapcsolatban lényegében semmilyen lehetséges elemzési szinten nincs konszenzus az ezzel foglalkozó kutatók között. Nem világos, hogy milyen fonetikai jellemzőkkel rendelkezik, illetve hogy rendelkezik-e egyáltalán állandó akusztikai-fonetikai jellemzőkkel. Kérdéses a fonológiai szerepe, és kapcsolata a morfológiai és szintaktikai jellemzőkkel. Az sem egyértelmű, hogy maga a hangsúly milyen elméleti keretben írható le a legmegfelelőbbben. Szintén kérdéses, hogy hogyan

1. A BESZÉD PROZÓDIAI JELLEMZŐINEK ÉSZLELÉSE

képzeltető el a hangsúly reprezentációja, a nyelvi jel pszichológia realitása és ennek idegrendszeri háttere (Fox, 2000).

A hangsúllyal kapcsolatban áttekintett, öt különböző területen alapuló szakirodalmi háttér az alábbiakban foglalható össze.

1. A *hangsúly akusztikai jellemzői* nem határozhatók meg egyértelműen. A főként az angol nyelvvel foglalkozó szakirodalomból az a konklúzió vonható le, hogy a hangsúly főként az f_0 és az intenzitás változásával jár együtt, de fontos lehet még a hossz is (Fry, 1958). Feltételezhető, hogy a hangsúlyos szótagok ezen akusztikai jellemzők közül mindig legalább az egyikben nagyobbak, mint a hangsúlytalan szótagok. Az akusztikai mérések eredményei ugyanakkor nem konzisztensek, és emiatt a hangsúlyt nem lehet megbízhatóan felhasználni az automatikus beszédfeldolgozó rendszerekben (Van Kuyk és Boves, 1999). A hangsúly produkciójára vonatkozó elektromiográfiás vizsgálatok ezzel szemben azt találták, hogy a beszédprodukciós rendszer szublingvális részéhez tartozó izmok szinte mindegyikénél nagyobb aktivitás mutatható ki a hangsúlyos szótag produkciójakor, szemben a hangsúlytalan szótaggal, és főként a bordaközi izmok aktivitása jár együtt igen megbízhatóan a hangsúly produkciójával (Fónagy, 1958).

2. A *hangsúly fonológiai hátterével* kapcsolatban feltételezhető, hogy az nem úgy írható le legjobban, mint egy adott fonológiai jegy jelenléte, vagy hiánya (\pm stress), és nem segíti a hangsúly megértését, ha egyszerűen annak szintjeit próbáljuk leírni (elsődleges, másodlagos, harmadlagos hangsúly, stb.). Ezzel szemben a hangsúly, mint *metrikus struktúra* leírása már sikeresebbnek mondható (Liberman és Prince, 1977). A metrikus fonológiai leírás a hangsúlyt relációs fogalmakban értelmezi, és ezzel megoldást kínál a hangsúly variábilis akusztikai jellemzőinek kezelésére: nem a hangsúly fizikai megjelenése a lényeg, hanem hogy a hangsúlyos és a hangsúlytalan szótag eltérjen egymástól. A metrikus fonológiai azt állítja, hogy a beszédben léteznek észrevehető lüktetések, amelyek közül bizonyosok kiemelkednek a többi közül, és ez a kiemelkedés egy szabályos és hierarchikus rendszert alkot. Az, hogy pontosan milyen szintek jönnek létre, nyelvenként eltérő. A magyarban a hangsúly rendszer viszonylag egyszerű volta miatt feltételezhetően elegendő egy a hangsúlyos és hangsúlytalan szótagokat elkülönítő metrikus szint, és egy a frázisok szintjén értelmezett intonációs szint.

3. A *hangsúly produkciójával* kapcsolatban azt mondhatjuk, hogy a produkciós folyamatban feltételezhetően létezik egy olyan szakasz, amikor az egymástól függetlenül reprezentált szegmentális és szupraszegmentális struktúra összekapcsolása megtörténik (Levelt, 2003). A szavak produkciójához

nélkülözhetetlen az, hogy a produkciós folyamatban valamikor rendelkezésre álljon egy olyan reprezentáció, amely nem csak a szavakat alkotó fonémákat jelöli ki, hanem a szavak szótag szerkezetét, és azok hangsúlymintázatát is. Ugyanakkor nem világos, hogy a hangsúlymintázat milyen módon reprezentálódik: ez minden egyes szó esetében rendelkezésre áll a mentális lexikonban, vagy pedig komputációs folyamatok révén, a fonológiai szabályok segítségével hozzuk létre a produkciós folyamat során. Feltételezhető, hogy mindkettő egyszerre létezik, és hogy nyelvenként eltér, hogy melyik lényegesebb (Cutler és Isard, 1980). A magyarban a hangsúly szabályos jellege miatt feltételezhetjük, hogy egy pusztán szabály alapú rendszer is megfelelően képes a hangsúlymintázat kijelölésére.

4. A *hangsúlynak a beszédfeldolgozásában betöltött funkciójával* kapcsolatban azt mondhatjuk, hogy elsősorban a lexikai feldolgozásban lehet ennek szerepe. Egyes elméletek szerint a lexikai feldolgozási folyamatban a hangsúlyos szótag lehet a beszédfolyam azon kiemelkedő eleme, amely lehetővé teszi a folyamatos beszéd szavakra történő szegmentációját, és ezáltal a lexikai keresési folyamat beindítását (Cutler és Norris, 1988). Ugyanakkor léteznek olyan elképzelések, amelyek szerint a hangsúly csak egy a lexikai feldolgozás során felhasznált kulcsok közül, és elsősorban akkor vesszük igénybe, amikor a hallási környezet nem optimális, például zajos környezetben (Mattys és munkatársai, 2005). Természetesen a szegmentációs kulcsok esetében is feltételezhető, hogy az egyes nyelvek eltérő módon alkalmazzák ezeket. Így a magyarban a szabályosság miatt a hangsúly szerepe nagyobb lehet, mint például az angolban.

5. A *hangsúlyfeldolgozás illetve produkció idegrendszeri háttere* kapcsán az afáziás betegekkel folytatott vizsgálatok azt mutatják, hogy a hangsúly és a szegmentális információ feldolgozása disszociálódhat (Baum és Pell, 1999). A hangsúly feldolgozás és produkció zavara elsősorban a bal féltekei sérülésekhez kapcsolódik. Hasonló eredményeket mutattak ki a hangsúly feldolgozás funkcionális neuroanatómiai hátterét tanulmányozó vizsgálatok, amennyiben a hangsúlymintázat feldolgozása során a bal oldali szuperior temporális területék aktivációját találták.

Disszertáción központi kérdése mindezen szakirodalmi adatok fényében az volt, hogy mit tudhatunk meg az agyi elektromos aktivitás változásainak vizsgálata révén a hangsúly idegrendszeri reprezentációjáról. Alapvető kérdésem arra vonatkozott, hogy a hangsúly, hasonlóan a szegmentális jellemzőkhöz, feldolgozható-e automatikus, azaz preattentív módon. Arra is kerestem a választ, hogy a hangsúlymintázat agyi feldolgozása mennyiben támaszko-

1. A BESZÉD PROZÓDIAI JELLEMZŐINEK ÉSZLELÉSE

dik a hangsúllyal kapcsolatos fizikai jellemzők feldolgozására, illetve mennyiben játszanak szerepet ebben a hosszú távú reprezentációk. Valamint az is érdekelt, hogy a hangsúlymintázat hosszú távú reprezentációja a mentális lexikonban specifikált leírás, vagy általános szabály formájában működik-e.

Az áttekintett elméletek és empirikus adatok alapján azt feltételezhetjük, hogy a hangsúlymintázat feldolgozása nem csak adatvezérelt, azaz bottom-up módon történik, hanem a top-down reprezentációk is szerepet játszanak a feldolgozási folyamatokban. A hangsúlymintázat feldolgozása feltehetőleg a beszédhangok feldolgozásával analóg módon zajlik. A beszédhangok feldolgozása kapcsán a Näätänen-féle modellt (Näätänen, 2001) szerint a beszédhangok feldolgozása a hosszú távú fonéma trace-eken, mint felismerési sablonokon alapul. Ennek alapján azt feltételeztük, hogy a prozódiai információ feldolgozása során is rendelkezésre állnak hosszú távú reprezentációk, amelyeket *prozódiai template*-eknek, vagy specifikusan a hangsúly kapcsán *hangsúly template*-eknek nevezhetünk. A template elnevezést azért tartjuk jobbnak a trace-nél a hangsúlyra vonatkoztatva, mert úgy véljük, hogy a hangsúly esetében nem egyszerűen emlékezeti nyomokról van szó, hanem mintázatokat tároló absztrakt reprezentációkról. Az áttekintett szakirodalmi adatok alapján a hangsúly template-eket úgy képzeljük el, hogy azok a szavak hangsúlymintázatát a metrikus szerkezetnek megfelelően tárolják, azaz a szótagokhoz kapcsolt hierarchikus hangsúlyszintek formájában, ahogy azt a Levett-modell (Levelt, 2003), illetve a Hayes-féle elképzelés (Hayes, 1995) leírja.

A hangsúlymintázat feldolgozásának EKP korrelátumai kapcsán úgy véltük, hogy a fonéma trace-ek működéséhez hasonlóan amennyiben a feldolgozott hangsúlymintázat eltér a reprezentációban meghatározott mintázattól, akkor aktiválódnak a változásdetekiós mechanizmusok, és megjelenik az EN komponens.

Mindezek alapján az agyi elektromos válaszokkal követni kívánt hangsúlyfeldolgozásra vonatkozóan az alábbi predikcióink lehetnek. Ha a hangsúlymintázat feldolgozásában top-down információk is szerepet játszanak a bottom-up információkon kívül, akkor a beszéd- és az ugyanolyan akusztikai jellemzőkkel rendelkező, de nem-beszéd jellegű ingerek feldolgozása során eltérő agyi válaszokat fogunk regisztrálni. Ha a top-down információ a mentális lexikonból származik, akkor a lexikonban tárolt jelentéssel rendelkező szavak, illetve a lexikonban reprezentációval nem rendelkező értelmetlen állszavak feldolgozása eltérő agyi válaszokkal fog együtt járni. Ha a top-down információ általános, szabály alapú template jellegű, akkor viszont a jelentés-

sel rendelkező szavak, és a jelentéssel nem rendelkező, de szó szerű álszavak feldolgozása hasonló agyi válaszokat fog kiváltani.

1.3. Tézisek

A fentiek alapján disszertációm tézisei az alábbiak voltak.

1. A hangsúly feldolgozása a szegmentális jellemzőkhöz hasonlóan preattentív módon valósul meg. A hangsúlymintázat sértésének feldolgozása EN komponenst vált ki.
2. A hangsúly a szegmentális jellemzőkhöz hasonlóan rendelkezik hosszú távú beszédspecifikus reprezentációval. Ezért a beszéd jellegű, és az ugyanolyan akusztikai jellemzőkkel rendelkező, de nem beszéd jellegű információk feldolgozása eltérő agyi válaszokat eredményez.
3. A hangsúly reprezentációja egy általános és absztrakt szabály, a hangsúly template formájában létezik, és a prelexikális szinten van jelen. Emiatt a szavak és álszavak feldolgozása során egyaránt aktív. A hangsúlymintázat feldolgozása során a nyelvi inputot összevetjük a hangsúly template-tel, és ha a nyelvi input nem felel meg a hangsúly template-nek, akkor az kiváltja az EN komponenst.

Mindezen kérdések megválaszolására öt eseményhez kötött agyi potenciál kísérletet végeztünk el, amelyek mindegyikében passzív kakukktojás paradigmákban vizsgáltuk az agyi elektromos aktivitás változását a különböző ingerparaméterek hatására.

1.4. Eredmények

1.4.1. 1. kísérlet

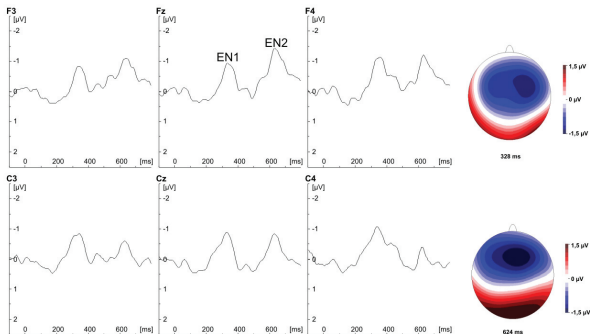
Az 1. kísérletben az volt a célunk, hogy megállapítsuk, képes-e az emberi agy a hangsúlymintázatban bekövetkező változás észlelésére preattentív helyzetben. Ezen kívül a jól ismert beszédhang kontrasztok feldolgozását is vizsgáltuk, annak érdekében, hogy a szegmentális és szupraszegmentális információk feldolgozását össze tudjuk hasonlítani. A beszédhang kontrasztok feldolgozása esetében azt vártuk, hogy a zöngésségi időben eltérő szókezdő

1. A BESZÉD PROZÓDIAI JELLEMZŐINEK ÉSZLELÉSE

beszédhangok automatikus feldolgozása kiváltja a jól dokumentált Eltérési Negativitás (EN) komponenst. A hangsúlymintázat feldolgozása kapcsán az volt az elvárásunk, hogy a magyar nyelvnek megfelelő hangsúlymintázat (a szó első szótagja hangsúlyos) sértése esetében (a szó második szótagja hangsúlyos) a nyilvánvaló eltérés, vagyis a második szótagon megjelenő hangsúly detektálása kiváltja az EN-t.

A kísérletben három különböző ingerszót alkalmaztunk, egy standard ingert és két abból származtatott deviáns ingert. A standard egy kétszótagú, értelmes szó volt („banán”). A két eltérő inger egyike, az úgynevezett „fonéma deviáns” abban tért el a standardtól, hogy a szókezdő „b” hang helyett annak zöngétlen párja, „p” szerepelt (vagyis a „panán” szót kaptuk). A másik úgynevezett „hangsúly deviáns” szupraszegmentális struktúrájában tért el a standardtól: a szabályos magyar kiejtéstől eltérően ennél a szónál a második szótagon volt a hangsúly.

Eredményeink azt mutatják, hogy mind a fonéma, mind a hangsúly deviáns inger szignifikáns EN-t váltott ki a fronto-centrális elektródákon, vagyis mind a szegmentális, mind a szupraszegmentális eltérés preattentív detekciója megtörtént. A hangsúly deviáns által kiváltott EKP-ok (ld. 1.1 ábra) azt mutatják, az a fonéma deviánstól eltérő EKP-okat vált ki, mégpedig két egymást követő EN-t, a fonéma deviáns által kiváltott egy helyett.



1.1. ábra. Az 1. kísérletben a hangsúly deviánsra adott különbségi görbe.

Ez arra utal, hogy a szegmentális és szupraszegmentális jellemzők feldolgozása eltérő módon, és egymástól függetlenül történik. A kérdés természetesen az, hogy a két egymást követő EN hogyan interpretálható. A hangsúlymintázat megváltozása ugyanis két különböző szinten is értelmezhető. Egyrészt a standard és deviáns inger jelentősen eltér egymástól az akusztikai jellemzőikben. Másrészt a hangsúlymintázat megváltozása az akusztikai változáson kívül fonológiai jellegű változással is együtt jár, amennyiben egy az első szótagján hangsúlyos szó helyett egy olyan szót hallunk, amely a második szótagján hangsúlyos. Ezen értelmezés szerint tehát a két EN a deviáns szó két szótagján található fonológiai jellegű eltérések detekcióját jelzi: az első szótagon mutatkozó hangsúly hiányt, és a második szótagon lévő plusz hangsúlyt.

1.4.2. 2-4. kísérlet

A 2-4. kísérletben azt vizsgáltuk, hogy az első kísérletben kapott elektrofiziológiai eredmények (az egymást követő két EN megjelenése) mennyiben tulajdoníthatóak annak, hogy itt elsősorban az ingerek akusztikai jellegzetességei kerülnek feldolgozásra. Ha az első kísérletben robusztus módon megjelent EN-ek csak az ingerek közötti akusztikai különbséget tükrözik, akkor hasonló módon kell megjeleníteniük olyan nem-beszéd jellegű ingerek esetében is, amelyek tartalmazzák az eredeti nyelvi ingerek akusztikai jellemzőit. Ha azonban a feldolgozás az ingerek nyelvi jellemzőire támaszkodik, és az EN-nel kísért eltérés detekciója ennek alapján történik, akkor a nem-beszéd ingerek esetében nem jelenik meg hasonló EKP mintázat.

A hangsúly akusztikai jellemzőit három különböző helyzetben modelleztük: alul-szűrt ingerekkel (2. kísérlet), csak az intenzitás változásában eltérő ingerekkel (3. kísérlet), és csak a hangmagasság változásában különböző ingerekkel (4. kísérlet). Az alul-szűrt ingerek alkalmazása bevett módszer a prozódiai jellemzők feldolgozását vizsgáló kísérletekben (ld. pl. Behrens, 1985), mivel ezzel a módszerrel olyan ingerek hozhatók létre, amelyek a szűrés miatt elveszítik beszéd jellegüket és érthetőségüket, ugyanakkor a prozódiai jellemzőket megvalósító akusztikai paraméterek az eredeti ingerekhez hasonlóak maradnak. A két másik modellben egy-egy olyan akusztikai paramétert választottunk ki (intenzitás és f_0), amelyekről feltételezhető volt, hogy fontos szerepet játszanak a hangsúly feldolgozásában. Ezek feldolgozását külön-külön vizsgáltuk meg úgy, hogy egy nem-beszéd jellegű ingert hoztunk létre, amely vagy az egyik, vagy a másik paraméternek az eredeti ingerekhez ha-

1. A BESZÉD PROZÓDIAI JELLEMZŐINEK ÉSZLELÉSE

sonló változását tartalmazta.

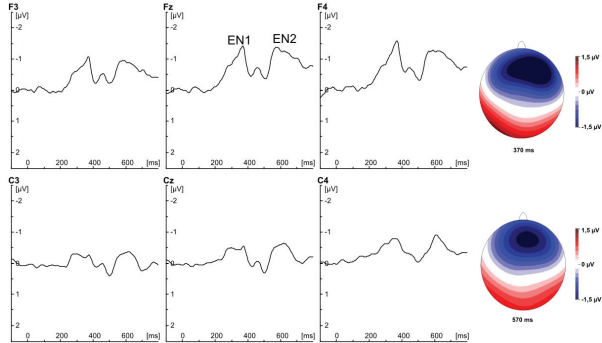
Összefoglalva a 2-4. kísérletek eredményeit, azt az eredményt kaptuk, hogy azok a nem-beszéd jellegű ingerek, amelyek a hangsúlymintázat változásának különböző akusztikai elemeit ragadták meg, azonosan eltérő EKP mintázatot eredményeztek a beszéd jellegű ingerekkel kiváltott EKP mintázattal történő összehasonlításban. Az eltérés lényege az volt, hogy egyik nem-beszéd jellegű inger sem váltotta ki azt az EN-t, amely beszédingerek esetében az inger kezdetéhez kapcsolódott, miközben a háromból két esetben kiváltható volt az EN2.

Ezek alapján kijelenthetjük, hogy a hangsúlymintázat eltéréseivel kiváltott EKP-ok változása csak részben tulajdonítható annak, hogy az ingerek akusztikai jellemzői kerülnek diszkriminatív feldolgozásra. Természetesen továbbra is kérdés, hogy pontosan miként interpretálható a beszédingerek kapcsán megjelenő két EN. A négy kísérlet eredményei mindenesetre azt jelzik, hogy az EN2 elsősorban a kiemelkedő (szátiens) akusztikai információ feldolgozásához kapcsolódik, és a nem-beszéd jellegű ingerek feldolgozásakor is megjelenik. Úgy tűnik, azonban hogy a kiinduló kísérletben regisztrált EN1 komponens beszédspecifikus, amennyiben nem jelenik meg nem-beszéd jellegű ingerek esetén.

1.4.3. 5. kísérlet

Az 5. kísérletben egy értelmetlen álszó esetében vizsgáltuk a hangsúlymintázat sértésének feldolgozását. A kísérlet első részében az előző kísérletekhez hasonlóan egy passzív kakukktójas paradigmában a szabályos hangsúlymintázatú álszó szerepelt standardként, és a szabálytalan hangsúlyozású (hangsúly a második szótagon) deviánsként. A kísérlet második részében megcseréltük az ingereket, és a szabálytalan hangsúlyozású álszó lett a standard, a szabályos pedig a deviáns. A kísérletben külön figyelmet fordítottunk arra, hogy az ingerszó mindkét szótagja azonos legyen, és a két szótag valóban csak a hangsúlymintázatban térjen el egymástól. Hipotézisünk az volt, hogy az álszavak esetében a hangsúlymintázat sértése hasonló feldolgozási mintázatot eredményez, mint a szavak esetében. Az eredmények valóban ezt mutatták, amennyiben 300 és 600 ms körül megjelent egy-egy EN, és ez szinte az összes vizsgált elektródán szignifikáns volt (ld. 1.2 ábra).

Ez az eredmény tehát arra utal, hogy a hangsúlymintázat feldolgozása a szó lexikai reprezentációjától függetlenül történik meg, és szavak és álszavak esetében hasonló módon alakul. Ugyanakkor, ahogyan a 2-4. kísérletek ered-



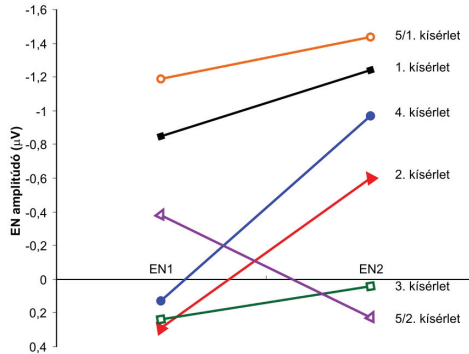
1.2. ábra. Az 5. kísérletben a deviánsra adott különbségi görbe.

ményei mutatták, nem-beszéd jellegű ingerek esetében egy eltérő feldolgozási mintázatot kapunk, ami tehát arra utal, hogy a jelen kísérletekben kapott dupla EN specifikusan a beszéd jellegű ingerek feldolgozását kíséri.

Az ötödik kísérletben alkalmazott fordított feltétellel annak bizonyítása volt a célunk, hogy a szabályos és szabálytalan hangsúlyozású ingerek a standard és deviáns helyzetben eltérő módon viselkednek. Azt vártuk, hogy a szabálytalan hangsúlyozású inger nem alkalmas arra, hogy egy megfelelő, a deviánsához való hozzámérést lehetővé tevő nyomot hozzon létre, mivel önmagában is sérti a feltételezett hosszú távú reprezentációt, és emiatt nem fog kiváltódni az EN. A kapott eredmények valóban ezt mutatták: sem az első EN sem a második EN nem volt szignifikáns.

Az eredmények illusztrálására az 1.3 ábra összefoglalja a 6 kísérlet fő eredményeit. A grafikon jól szemlélteti azt, hogy az EN1 komponens csak azokban a kísérletekben jelent meg, amelyekben beszéd jellegű ingereket alkalmaztunk, az EN2 viszont lényegében minden helyzetben megjelent. Illetve az 5. kísérlet 2. feltétele egyik kísérlethez sem hasonlító eredményt hozott, mivel az EN2 mindegyik kísérletben talátnál kisebb volt, az EN1 viszont a beszéd és nem-beszéd ingereket alkalmazó kísérletek között helyezkedett el.

1. A BESZÉD PROZÓDIAI JELLEMZŐINEK ÉSZLELÉSE



1.3. ábra. Az 1-5. kísérlet eredményeinek összefoglalása. Az ábrán az EN1 és EN2 komponensek csúcsamplitúdóinak átlaga látható a 6 kísérleti helyzetben.

1.5. Megbeszélés

Eredményeink alapján tehát azt feltételezzük, hogy a hangsúlymintázat sértésének feldolgozása legjobban egy olyan elméleti keretben értelmezhető, amely figyelembe veszi, hogy a hangsúly absztrakt fonológiai template-jei hosszú távú reprezentációként állnak a nyelvi feldolgozó rendszer rendelkezésére, s ezek a valamennyi legális szóalakra (szavak, álszavak) vonatkozó általános hangsúlyozási jellegzetességeket tartalmazzák. A preattentív feldolgozás során ezek a template-ek szolgálnak a változásdetekció alapjául, és az ezektől való eltérés detekciója eredményezi az Eltérési Negativitás megjelenését.

A hangsúly template szerkezete úgy képzelhető el, hogy az tartalmazza a szavak szótagszerkezetét, ezek metrikus lábakra történő csoportosítását, és azt, hogy a szóhangsúly a metrikus láb mely részére esik. A magyar nyelvben eszerint a template azt írja le, hogy a szavakat alkotó szótagok úgynevezett szillabikus trocheusokba (a Hayes, 1995 által feltételezett láb típusok közül az egyik) szerveződnek, és a szókezdő szótag rendelkezik a hangsúllyal.

A hangsúly template-ek szerepét a beszéd percepciójában a következőképpen képzelhetjük el. Az akusztikai input feldolgozása során elsőként kinyerjük

a szegmentális jellemzőket (fonémák), illetve létrehozuk a szótag szerkezetet. Ezt követően meg kell állapítani, hogy a szótagok hangsúlyosak vagy hangsúlytalanok. Feltételezhetően már ebben a folyamatban lényegesek lehetnek a hangsúly template-ek, mivel felismerési sablonokként működve lehetővé teszik az egyébként variábilis akusztikai jellemzőkkel rendelkező hangsúly feldolgozását. Azaz a feldolgozás során a hangsúly template-ek aktiválódása jelzi a szótagok hangsúlyos vagy hangsúlytalan voltának felismerését, vagyis a kategorizációt.

Itt tehát feltételezünk a szófeldolgozási folyamatban egy olyan reprezentációs szintet, amely a szótagok hangsúlyos-hangsúlytalan leírását tartalmazza. Ez megfelel a Grosjean és Gee (1987) által a szófeldolgozás folyamatában feltételezett közbenső (intermediary) reprezentációnak. Grosjean és Gee (1987) elmélete szerint a beszédfeldolgozás során az akusztikai input egy olyan közbenső reprezentációba íródik át, amely a fonetikai szegmensek hangsúlyos-hangsúlytalan leírását tartalmazza. A lexikai hozzáférés pedig ezen közbenső reprezentáció alapján indul el, mégpedig a hangsúlyos szótagoktól kezdődően, amelyek a feltételezhető szókezdeteket jelölik.

Ennek alapján a kísérleteinkben kapott EKP eredmények értelmezése véleményünk szerint az, hogy a feldolgozás során, amennyiben nyelvi ingerről van szó, a feldolgozó rendszer az inputot nem csak az előzőleg beérkezett memórianyomokkal veti össze, hanem figyelembe veszi azt is, hogy az ingerek mennyire illeszkednek a hangsúly template-hez. A hangsúly deviáns esetében a szó első szótagja nem felel meg a template-ben specifikált jellemzőknek, emiatt az első szótagot a rendszer nem képes hangsúlyosnak kategorizálni. Ennek következtében az összemérés során eltérést jelez a feldolgozó rendszer, EKP korrelátumaként pedig EN regisztrálható. Nem kíséri EN megjelenése az összemérést, ha nem-beszéd ingereket dolgozunk fel, hiszen ekkor nem aktiválódnak az összemérés alapjául szolgáló template-ek, illetve ehhez hasonlóan szintén nem jelenik meg EN akkor, ha a deviáns inger nem sérti magát a template-et.

Saját eredményeink a hangsúly template-ekkel kapcsolatban tehát azt mutatják, hogy ezek feltételezhetően irányítják a hangsúllyal kapcsolatos feldolgozási folyamatokat, és arra utalnak, hogy a szófeldolgozás során a szavakat minden esetben automatikus módon hozzámérjük a template-ekhez. Az EKP eredmények azt is mutatják, hogy ez a feldolgozás beszédspecifikus jellegű, de szublexikálisan működő folyamat. Véleményünk szerint a kísérleteink során kapott elektrofiziológiai eredmények a nyelvfeldolgozási folyamatoknak egy olyan szakaszába engednek betekintést, amelynek során a nyelvi inputból

1. A BESZÉD PROZÓDIAI JELLEMZŐINEK ÉSZLELÉSE

az észlelőrendszer kinyeri az adott nyelvre jellemző, specifikusan szupraszegmentális jellegű információkat, és ezekből felépíti azokat a reprezentációkat, amelyek a nyelvi inger fonológiai leképezését teszik lehetővé. Ilyen értelemben eredményeink empirikus bizonyítékkal szolgálnak amellett, hogy a fonológiai elméletek által feltételezett hangsúly reprezentációk valós agyi konstruktumok, amelyek működése megragadható az agyi elektromos változások online tanulmányozása révén.

1.6. A disszertáció témájában megjelent publikációk

Honbolygó, F., Csépe, V. (2009). MMN to stress change in speech: rule or regularity? *Frontiers in Human Neuroscience. Conference Abstract: MMN 09 Fifth Conference on Mismatch Negativity (MMN) and its Clinical and Scientific Applications.* (konferencia absztrakt)

Honbolygó, F., Csépe, V. (2008) The processing of suprasegmental cues is speech specific as evidenced by ERPS. *International Journal of Psychophysiology*, 69 (3), 238-238. (konferencia absztrakt)

Honbolygó, F., Csépe, V. (2007) A beszéd szupraszegmentális jellemzőinek automatikus észlelése: eseményhez kötött agyi potenciál vizsgálat. *Pszichológia*, 27, 71-85.

Honbolygó, F., Csépe, V., Fekésházy, A., Emri, M., Márián, T., Sárközy, G. and Kálmánchey, R. (2006) Converging evidences on language impairment in Landau-Kleffner Syndrome revealed by behavioral and brain activity measures: A case study. *Clinical Neurophysiology*, 117/2, 295-305. Impakt faktor: 2.540

Honbolygó, F., Csépe, V., Ragó, A. (2004) Suprasegmental Speech Cues are Automatically Processed by the Human Brain: a Mismatch-Negativity Study. *Neuroscience Letters*, 363/1. 84-88. Impakt faktor: 2.020

1.6 A disszertáció témájában megjelent publikációk

2.

The perception of prosodic features of speech

2.1. Introduction

In my dissertation I studied the neural basis of speech perception. Speech perception involves several different processes in the brain, which make it possible to retrieve the linguistic message from the acoustic input. Of all these processes, my dissertation covers mainly those related to the prosodic part of speech, and especially to stress. The study of the prosodic features of speech received little attention in speech perception research, in spite of the fact that prosodic features are as much part of the linguistic input as the segmental features. Therefore, the main goal of my dissertation is to draw attention to the importance of prosodic features.

The key subject of my dissertation is the investigation of electrical brain activity elicited by speech processing. The stimulus related synchronized change of EEG signal recorded from the human head, i.e. the event-related brain potential (ERP) is probably the most successful method of cognitive neuroscience. Using ERPs it is possible to explore with relatively little effort those cognitive processes that would otherwise be difficult to investigate with the other methods of cognitive psychology. In my dissertation, I'd like to demonstrate that ERP components, with the proper experimental procedures, allow us to examine the functioning of representations proposed by linguistic or psycholinguistic models in a real-time, online way.

In the dissertation I employed a multidisciplinary approach for the study

of stress, because I believe that in order to understand this complex linguistic phenomenon, we have to take into account several theoretical frameworks. Following this approach, in the Literature review I present the linguistic, psychological and neuro-scientific theories of stress.

The theses of the dissertation were examined in five experiments. In all of the experiments, the method of event-related potentials was applied, and from the ERP components I focused on the Mismatch Negativity (MMN) component, as a brain electrical correlate of pre-attentive change detection.

2.2. Literature review

During the perception of speech, two kinds of acoustical information need to be processed: *segmental* information related to the speech sounds, and *suprasegmental* information, which is built on, spans over and modulates the speech sounds (Gósy, 2004).

The prosodical or suprasegmental features are usually contrasted with the segmental features. In the production of speech, the two kinds of features appear together as a complex acoustical signal, as they are produced simultaneously, and during the perception of speech, they have to be processed together as well. The segmental and suprasegmental features hold different functions. To simplify, we can say that segmental features contain information about the speech sounds, and suprasegmental features play delimitative functions, that is they separate linguistic units of different levels (words, phrases, sentences).

In the dissertation, I focus on one of the most debated of the prosodic features, that is stress. Concerning the nature of stress, there is no consensus on any of the possible analysis levels among the researches dealing with the topic. It is not clear what phonetical properties the stress has, or whether it has any stable acoustical-phonetical properties at all. It's phonological function, and its relation to morphological and syntactical features is questioned. It's accurate description in terms of theoretical frameworks is debatable. And it is unclear how stress is represented, what is the psychological reality of the linguistic signal, and what is its neural basis (Fox, 2000).

The literature reviewed about stress, based on five different theoretical background, can be summarized as follows.

1. The *acoustical properties* of stress can not be determined unambiguously. The literature dealing mostly with English language concludes that

2. THE PERCEPTION OF PROSODIC FEATURES OF SPEECH

stress is correlated with the change in f_0 and intensity, but length can be important as well (Fry, 1958). It seems that stressed syllables are higher in at least one of the above acoustical parameters, as compared to unstressed syllables. At the same time, acoustical measurements are inconsistent and because of that it is suggested that at present, stress can not be reliably incorporated into automatic speech recognition systems (Van Kuijk és Boves, 1999). Electromyographic measurement about the production of stress however found that in almost all of the muscles of the sub-laryngeal part of the speech production system show a higher electrical activity during the articulation of stressed syllable, as compared to unstressed syllables. Especially the activity of intercostal muscles seems to correlate well with the production of stress (Fónagy, 1958).

2. Concerning the *phonological background* of stress we can assume that it is not the best way to describe it as a presence or absence of a phonological feature ($[\pm \text{ stress}]$), and it does not help in its understanding if we specify simply its levels (primary, secondary, ternary stress, etc.). On the contrary, the description of stress as a *metrical structure* is a more successful approach (Liberman és Prince, 1977). The metrical phonological description analyzes stress in relational terms, and provides a solution to handle the variable acoustical properties of stress: these are not the physical properties of stress that are important, but the simple deviation of stressed and unstressed syllables. The theory of metrical phonology states that in the speech stream there are noticeable beats, some of which are more prominent than others, and the prominence creates a regular and hierarchical system. The exact form of the hierarchy differs from language to language. In Hungarian, because of the relatively simple stress system, one metrical level representing stressed and unstressed syllables, and one intonation level representing the phrase structure is presumably enough to describe the metrical hierarchy.

3. With regard of *stress production*, it is suggested that in the speech production process there is a stage where the independently represented segmental and suprasegmental structure is combined together (Levelt, 2003). For the production of words it is evidently required that somewhere in the process of speech production we have a representation that specifies not only the phonemes of the words, but also the syllable structure and the stress pattern of words. It is unclear however how the stress pattern is represented: it is available for each word in the mental lexicon, or it is created during the production process by computational mechanisms, based on phonological rules. It is suggested that both of these may be at work, but languages

differ with respect to which one is more important (Cutler és Isard, 1980). In Hungarian, because of the regular nature of stress, we can suppose that a single rule-based system is enough to specify the required stress pattern.

4. The function of stress in *speech perception* is proposed to be related to lexical processing. According to certain theories, the stressed syllable may be the prominent part of the speech stream that could be used to segment speech into words and therefore initiate lexical search (Cutler és Norris, 1988). Other theories suggest that stress is only one of the many possible segmentation cues in lexical processing, and it is mainly used when the listening environment is not optimal, e.g. in noisy environments (Mattys és munkatársai, 2005). Evidently, different languages can be supposed to employ different segmentation cues. Thus, because of its regularity, stress might play a more important role in Hungarian as compared to English.

5. Concerning the *neural basis* of stress perception and production, studies with aphasic patients show that the processing of stress and segmental information can dissociate (Baum és Pell, 1999). The deficit of stress perception and production is mainly related to left hemisphere lesions. Similar results were obtained from experiments investigating the functional neuro-anatomical basis of stress processing, that demonstrated the activation of left superior temporal areas during stress processing.

Based on the review of literature, the central question of my dissertation was that what can be revealed about the neural representation of stress by examining the changes in the brain's electrical activity. First, I dealt with the question whether stress can be processed in an automatic, pre-attentive way, similarly to segmental features. Second, I investigated to what extent the neural processing of stress patterns is based on the processing of physical parameters related to stress, and what is the possible role of long-term representations in this processing. Finally, I examined the nature of the long-term representations of stress, and the question here was if this representation is specified in the mental lexicon, or it is functioning as a general rule.

Taking into account the reviewed theories and empirical data, we assumed that the processing of stress patterns is not only data-driven, that is a bottom-up process, but top-down representations play an important role also. The processing of stress patterns is supposedly analogous to the processing of speech sounds. One theory about the processing of speech sounds (Näätänen, 2001) suggests that this processing is based on long-term phoneme traces, as recognition patterns. Correspondingly, we proposed that in

2. THE PERCEPTION OF PROSODIC FEATURES OF SPEECH

the processing of prosodic information there are long-term representations available, that we called prosodic template, or specifically for stress, stress templates. We preferred the term template to trace, because we think that in the case of stress the representations are not simple memory traces, but complex and abstract patterns. Derived from the literature, stress templates are presumed to store the stress pattern of words in the form of a metrical structure, that is as metrical levels associated with the syllables of words, as described by the Levelt-model (Levelt, 2003), or by the theory of Hayes (Hayes, 1995).

Concerning the ERP correlates of stress processing, we suggested that analogously to the way phoneme traces function, if the perceived stress pattern differ from the pattern defined in the stress template, then the change detection mechanism is triggered, and the MMN component is elicited.

To sum up, we formulated the following predictions about stress processing investigated by the measurement of brain electrical activity. If top-down information, in addition to bottom-up information play a role in the perception of stress patterns, then we can expect to record differing brain activities while processing speech, as compared to non-speech stimuli, having similar acoustical characteristics. If the top-down information originates from the mental lexicon, then meaningful words stored in the lexicon, and meaningless non-words having no lexical representation will elicit different brain responses. If however the top-down information is a general, rule-based template, then words and non-words will elicit similar brain responses.

2.3. Theses

Based on the arguments above, the theses of my dissertation are as follows.

1. Similarly to segmental features, the perception of stress is a pre-attentive process. The violation of stress pattern elicits the MMN component.
2. Similarly to segmental features, stress possesses a long-term, speech specific representation. Therefore the processing of speech, and non-speech information having similar acoustical features elicit differing brain responses.
3. The representation of stress is a general and abstract rule, functioning as a stress template at the pre-lexical level. Therefore the template is

active for the perception of both words and non-words. In the course of the processing of stress patterns, the linguistic input is compared to the stress template, and if the two do not match, the MMN component is elicited.

In order to investigate these question, we conducted five event-related brain potential experiments, in all of which we used a passive oddball paradigm to measure the changes in brain electrical activity following the different stimulus parameters.

2.4. Results

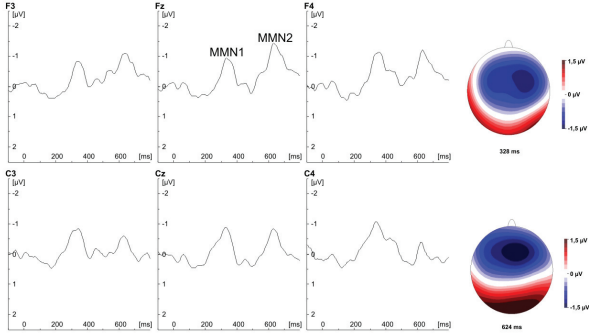
2.4.1. Experiment 1.

In Experiment 1. our goal was to find out if the human brain is capable of pre-attentively detect changes in the stress pattern. Furthermore, we investigated the well-known processing of speech sounds, in order to be able to compare the processing characteristics of segmental and suprasegmental information. In the case of speech contrasts, we expected to find the MMN component as a response to the automatic processing of voice-onset-time difference. In the case of stress processing, our expectation was that if the stress pattern corresponding to Hungarian (stress on the first syllable) is violated (stress on the second syllable), then the detection of this change elicits the MMN.

In the experiment we applied three different words, one as a standard, and two others as deviants. The standard was a two syllable long meaningful Hungarian word („banán”). The so-called phoneme deviant differed from the standard in its first phoneme, which was changed from „b” to „p” (i.e. we used to word „panán”). The second deviant, the so-called „stress deviant” differed from the standard in the suprasegmental structure, having the stress on the second, instead of the obligatory first syllable.

Our results showed that both the phoneme deviant and stress deviant elicited a significant MMN on the fronto-central electrode sites, which is interpreted as the pre-attentive change detection of both segmental and suprasegmental information. The ERPs elicited by the stress deviant (see figure 2.1) were however different from the ERPs elicited by the phoneme deviant, because for the stress deviant two consecutive MMN appeared, instead of only one.

2. THE PERCEPTION OF PROSODIC FEATURES OF SPEECH



2.1. figure. Difference curve to the stress deviant in Experiment 1.

This result confirms that the segmental and suprasegmental features are processed differently and probably independently. At the same time the interpretation of the two MMNs was unclear in the present experiment. The changes of stress patterns can be explained on at least two levels. First, the standard and stress deviant differed in their acoustical characteristics. Second, the change can be interpreted at the phonological level as well, as a change in the stressing of words. From Experiment 1., it was not possible to decide which interpretation is correct.

2.4.2. Experiment 2-4.

In order to be able to interpret the findings of Experiment 1., we conducted a series of experiments in which we manipulated the acoustical characteristics of stimuli. The aim of these experiments was to disprove that the two MMNs recorded in Experiment 1. as a response to the stress change was due only to the processing of the acoustical information related to stress change. We hypothesized that if in Experiment 1. the MMNs reflect only the acoustical change detection, then a similar MMN pattern should appear to non-speech stimuli having the same acoustical properties. But if the MMNs reflect language-related processing, then the MMN pattern elicited by non-speech stimuli should be different.

The acoustical characteristics of stress were modeled in three different situations: using low-pass filtered stimuli (Experiment 2.), with stimuli containing only intensity change (Experiment 3.) and using stimuli containing only pitch change (Experiment 4.). Low-pass filtering speech stimuli is a common experimental method in studies of prosodical processing, because it enables to create non-speech stimuli that nonetheless retain the acoustical properties important for prosody. In the two other models, we chose the acoustical properties (intensity and pitch) that were supposedly the most important features in stress processing. In both cases we created non-speech stimuli that kept the acoustical properties from the original stimulus in the given parameter, but there was no other acoustical change.

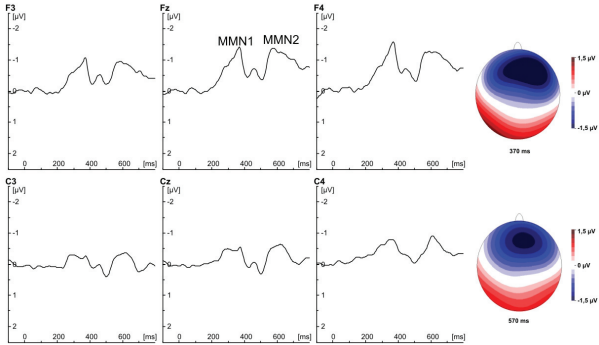
To summarize the results of Experiment 2-4., we obtained ERPs that were different as compared to those in Experiment 1. That is non-speech stimuli modeling the different characteristics of stress, and having the same acoustical properties as speech stimuli, did not elicit the same ERP pattern. The difference was such that non-speech stimuli elicited only the second MMN, but not the first. This result suggest that the MMNs elicited by the processing of stress pattern are attributable to other factors than processing only the acoustical differences, and presumably speech-specific processes are also playing an important role.

2.4.3. Experiment 5.

In Experiment 5. we investigated the processing of stress pattern violation in the case of a meaningless non-word. We applied two experimental conditions. In the first condition, the non-word with the regular stress pattern (stress on the first syllable) was the standard stimulus, similarly to Experiment 1. In the second condition however the non-word with the irregular stress pattern (stress on the second syllable) was the standard. In this experiment we wanted to assure that both syllables of the stimuli were the same, so that the stressed and unstressed syllables differed only in the stress pattern. We hypothesized that the violation of stress pattern elicits similar ERP responses in non-words as compared to words. The results confirmed this, as the deviant non-word stimulus elicited two MMNs, at around 300 and 600 ms (see figure 2.2).

This result suggests that the processing of stress patterns is independent of the lexical representation of a word, and similar ERP responses are elicited by words and non-words. At the same time Experiment 2-4. showed that the

2. THE PERCEPTION OF PROSODIC FEATURES OF SPEECH

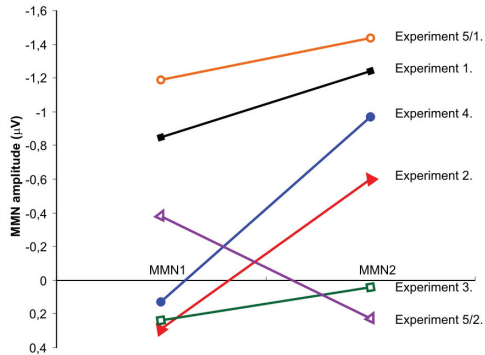


2.2. figure. Difference curve to stress deviant in Experiment 5.

processing of non-speech stimuli elicits different ERPs, indicating that the double MMNs are specific to the processing of speech information.

In the second condition of Experiment 5., where the non-word with the irregular stress pattern was the standard stimulus, our goal was to demonstrate that non-words with different stress patterns do not behave similarly in standard and deviant situations. We expected to find that the non-word with the irregular stress pattern is unable to form a trace that could serve as a basis for the comparison of standards and deviants, because the irregular standard violates the long-term representation in itself, and as a consequence, no MMN would be elicited. The results confirmed this, as none of the two MMNs appeared.

Figure 2.3 summarizes the main findings of the six experimental conditions. The figure illustrates nicely that MMN1 appeared only in those experiments where speech stimuli were applied, while MMN2 appeared in all conditions. In addition, the second condition of Experiment 5. showed very different results as compared to all other conditions, as the amplitude of MMN2 was smaller than in each of the other conditions, and the amplitude of MMN1 was in between of speech and non-speech stimuli.



2.3. figure. Summary of the results of experiment 1-5. The figure shows the peak amplitudes of MMN1 and MMN2 in the 6 conditions.

2.5. Discussion

Based on these results we suggest that the processing of stress pattern violation can be best interpreted in a theoretical framework that takes into account that the abstract phonological templates of stress are available for the language processing system as long-term representations, and these templates incorporate the general stress characteristics applied to all legal word forms (words and non-words). In the pre-attentive processing of stress, these template are the basis of change detection, and the detection of deviation from these templates results in the appearance of the MMN.

The inner structure of the stress template can be conceptualized as a representation including the syllable structure of words, the grouping of syllable into metrical feet, and the specification of the stressed syllable. Accordingly, in Hungarian the template describes that syllables are organized into the so-called syllabic trochees (one of the foot types hypothesized by Hayes, 1995), and the first syllable of the word should bear the stress.

The role of stress templates in the perception of speech can be conceived as follows. During the processing of acoustical input, the segmental features (phonemes) are extracted first, followed by the construction of syllables.

2. THE PERCEPTION OF PROSODIC FEATURES OF SPEECH

After that it is necessary to determine if the syllable are stressed or unstressed. We suggest that stress templates play an important role already here, because they might function as recognition patterns enabling the processing of the acoustically highly variable stress. That is in the course of processing, the activation of the stress template may signal the recognition of syllable as stressed or unstressed, in other words the categorization of syllables as stressed or unstressed.

We assume thus the existence of a representational level in speech processing that specifies the stressed or unstressed nature of syllables. This corresponds to the intermediary representation proposed by Grosjean és Gee (1987). According to the theory of Grosjean és Gee (1987), during speech perception the acoustical input is transformed into a representation describing the phonetical segments in terms of stressing. They suggest that the lexical access starts from these intermediary representations, as it is on this level that the stressed syllables, serving as specifications of the beginning of words, are first specified.

Taking into account all what is said, we think that the interpretation of our ERP results is that during the processing, in case of a linguistic stimuli the processing system compares the input not only with the previous memory traces, but also takes into consideration to what extent the input fits to the stress template. In the case of the stress deviant the first syllable of the word does not matches the characteristics specified in the template, and as a result the system can not categorize the first syllable as stressed one. Because of this, the system reports a deviation during the matching process and MMN can be registered, as an ERP correlate. No MMN appears during matching if non-speech stimuli are processed, because the templates serving as the basis for the matching are not activated, and similarly, no MMN appears when the deviant stimuli does not violate the template itself.

To summarize, our own results regarding the stress templates show that stress templates supposedly govern the stress perception processes. The results also indicate that the words are matched with the templates automatically during word processing. The ERP results point out that this processing is speech specific and sub-lexical. We think that the electrophysiological results obtained from our experiments allow us to gain insight into that part of language processing, where the cognitive system retrieves the specific suprasegmental information of a given language from the input, and from these information it builds up the representations enabling the phonological categorization of the linguistic stimuli. Thus we think that our results provide

empirical evidence that the stress representations assumed by the phonological theories are real neural constructs, the functioning of which can be captured by the online exploration of the brain electrical changes.

2.6. Publications related to the topic of the dissertation

Honbolygó, F., Csépe, V. (2009). MMN to stress change in speech: rule or regularity? *Frontiers in Human Neuroscience. Conference Abstract: MMN 09 Fifth Conference on Mismatch Negativity (MMN) and its Clinical and Scientific Applications.* (konferencia absztrakt)

Honbolygó, F., Csépe, V. (2008) The processing of suprasegmental cues is speech specific as evidenced by ERPS. *International Journal of Psychophysiology*, 69 (3), 238-238. (konferencia absztrakt)

Honbolygó, F., Csépe, V. (2007) A beszéd szupraszegmentális jellemzőinek automatikus észlelése: eseményhez kötött agyi potenciál vizsgálat. *Pszichológia*, 27, 71-85.

Honbolygó, F., Csépe, V., Fekesházy, A., Emri, M., Márián, T., Sárközy, G. and Kálmánchey, R. (2006) Converging evidences on language impairment in Landau-Kleffner Syndrome revealed by behavioral and brain activity measures: A case study. *Clinical Neurophysiology*, 117/2, 295-305. Impakt faktor: 2.540

Honbolygó, F., Csépe, V., Ragó, A. (2004) Suprasegmental Speech Cues are Automatically Processed by the Human Brain: a Mismatch-Negativity Study. *Neuroscience Letters*, 363/1. 84-88. Impakt faktor: 2.020

Irodalomjegyzék

- Baum, S. és Pell, M. (1999). The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology*, 13(8): 581–608.
- Behrens, S. (1985). The perception of stress and lateralization of prosody. *Brain and language*, 26(2): 332–48.
- Cutler, A. és Isard, S. (1980). The production of prosody. In Butterworth, B. (szerk.), *Language Production*, 245–269. Academic Press, London.
- Cutler, A. és Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1): 113–121.
- Fónagy, I. (1958). A hangsúlyról. *Nyelvtudományi Értekezések*, 18.
- Fox, A. (2000). *Prosodic features and prosodic structure: the phonology of suprasegmentals*. Oxford University Press, USA.
- Fry, D. (1958). Experiments in the perception of stress. *Language and Speech*, 1(2): 126–152.
- Gósy, M. (2004). *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Grosjean, F. és Gee, J. P. (1987). Prosodic structure and spoken-word recognition. *Cognition*, 25: 135–155.
- Hayes, B. (1995). *Metrical stress theory: principles and case studies*. University of Chicago Press, Chicago.
- Levett, W. (2003). Producing spoken language: A blueprint of the speaker. In Brown, C. és Hagoort, P. (szerk.), *The Neurocognition of Language*, 83–122. Oxford University Press, Oxford.

- Liberman, M. és Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8(2): 249–336.
- Mattys, S. L., White, L., és Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General*, 134(4): 477–500.
- Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, 38(01): 1–21.
- Van Kuijk, D. és Boves, L. (1999). Acoustic characteristics of lexical stress in continuous telephone speech. *Speech Communication*, 27(2): 95–112.